Bull Environ Contam Toxicol (2010) 85:109–115
DOI 10.1007/s00128-010-0065-2

QSAR Models for Predicting Toxicity of Polychlorinated Dibenzo-*p*-dioxins and Dibenzofurans Using Quantum Chemical Descriptors

Jianxiong Diao · Yang Li · Shuqiong Shi · Ye Sun · Ying Sun

Received: 5 May 2010 / Accepted: 2 July 2010 / Published online: 14 July 2010

Abstract By partial least square regression, simple quantitative structure–activity relationship (QSAR) models were developed for the toxicity of polychlorinated dibenzo-*p*-dioxins and dibenzofurans (PCDD/Fs). Quantum chemical descriptors computed by semi-empirical PM3 method were used as predictor variables. Three optimal QSAR models are developed for 25 PCDDs, 35 PCDFs, 25 PCDDs and 35 PCDFs together, respectively. The cross-validated Q^2_{cum} values for the three QSAR models of 25 PCDDs, 35 PCDFs, 25 PCDDs and 35 PCDFs together are 0.816, 0.629 and 0.603, respectively, indicating good predictive capabilities for the biological toxicity of these PCDD/Fs. The present study suggests that quantum chemical descriptors of POPs indeed govern the binding affinity of these chemicals for aryl hydrocarbon receptors. Moreover, different models contain different molecular descriptors to define respective equation, which suggests that the relationship between molecular structure and the binding affinity of these chemicals for aryl hydrocarbon receptors is complex.

Keywords PCDD/Fs · Quantum chemical descriptors · QSARs · Toxicity

J. Diao · Y. Li · Y. Sun · Y. Sun (✉)
Department of Chemistry, China Agricultural University, No. 2 Yuanmingyuan Xi Road, Haidian District, 100193 Beijing, People's Republic of China
e-mail: sunying@cau.edu.cn

S. Shi
State Key Laboratory of Water Environment Simulation, School of Environment, Beijing Normal University, 100875 Beijing, People's Republic of China

Polychlorinated dibenzo-*p*-dioxins and dibenzofurans (PCDD/Fs) are ubiquitous contaminants that persist in the environment. As one class of typical persistent organic pollutants (POPs), they can bioaccumulate through the food chain, and pose a risk of causing adverse effects to human health. PCDD/Fs have been detected in many kinds of environmental matrices even in remote areas such as the Arctic (Bjorn and Krister 2000; Niu et al. 2003). Because of their properties, PCDD/Fs have raised concern about their adverse effects such as carcinogenicity, teratogenicity, and mutagenicity on organisms and human (Bayen et al. 2007; Bustnes et al. 2006). In many epidemiologic researches, it was concluded that human in a risk of PCDD/Fs exposure are liable to suffer from certain diseases (Huang et al. 2006). Therefore, investigations on toxicity of PCDD/Fs are of great importance to understand their risk to human health.

It is significant to make use of the toxicity data of PCDD/Fs to evaluate their risk to organisms and further adopt effective measures to reduce their adverse effects. However, because of high cost, time-consuming process, limits of detection and lack of adequate standard materials, toxicity data are rather scarce for nongenotoxic adverse effects of PCDD/Fs. In order to conquer these problems and quickly estimate the environmental behaviors of PCDD/Fs, quantitative structure–activity relationship (QSAR) models, which correlate and predict toxicity data of PCDD/Fs from their molecular structural descriptors, provide valuable approach in research into the toxicity of PCDD/Fs without any experiments, were widely applied to evaluate and predict toxicity of PCDD/Fs efficiently (Netzeva and Schultz 2005; Juranić et al. 2006; Devillers et al. 2006).

QSAR models were developed to study the toxicity of PCDFs by Clare (2006), who found that the combination of

the Pearson-Parr softness and the orientation of nodes in the occupied π orbitals together accounts for up to 95% of the ability of chemicals binding to the aryl hydrocarbon receptor. Ashek et al. (2006) performed QSAR research on the toxicity of dioxins and dioxin-like compounds. They suggested a model comprises of four structurally different compounds, which offers a good predictive ability for various ligands. Previous studies have shown that reliable QSAR models are not only applied to predict toxicity and provide basic data to risk assessment, but also used to explain the toxicity mechanisms.

Quantum chemical descriptors can be easily obtained by computation, can clearly describe defined molecular properties, and are not restricted to closely related compounds. Consequently, reliable QSAR models developed using quantum chemical descriptors have been applied in various domains, such as photodegradation of PCDD/Fs adsorbed on spruce needles under sunlight irradiation (Niu et al. 2005), binding affinity between phenolic chemicals and estrogen receptors (Hu and Aizawa 2003), *n*-octanol/water partition effects of PCBs (Padmanabhan et al. 2006), and the toxicity of hydroxylated and quinoid PCB metabolites (Niu et al. 2007). A question is thus posed: is it possible to develop reliable QSAR models for PCDD/F toxicity using quantum chemical descriptors? The answer to this question is of importance to investigate the applicable domains of quantum chemical descriptors on developing QSAR models.

As partial least square algorithm can analyze data with strongly collinear, noisy and numerous *X* variables (Wold et al. 2001), it can be used to develop reliable QSAR models to predict the bioactivity of a series of organic chemicals (Long and Niu 2007; Niu et al. 2007; Zhang et al. 2008). It is not only searching the relationship between a matrix *Y* (containing dependent variables) and a matrix *X* (containing predictor variables), but also reducing the dimension of the matrices while concurrently maximizing the relationship between the descriptors. The purpose of this study is to develop reliable QSAR models for the binding affinity of PCDD/Fs for aryl hydrocarbon receptors using quantum chemical descriptors and depending on partial least square algorithm.

Materials and Methods

The experimental toxicity values, i.e. binding affinity values (log $1/EC_{50}$) of PCDD/Fs and dioxin-like compounds for aryl hydrocarbon receptors in vitro rat hepatocyte assays reported, were adopted in this study (Cheung and McKinney 1989; Okey 1990; Waller and McKinney 1992). These toxicity values were collected by Ashek et al. (2006). The compounds and their respective biological activity data used in this QSAR analysis are represented in Table 1. The results show that an approximately threefold difference exists between the known log $1/EC_{50}$ value for dibenzofuran (3.00) and that for 2,8-dibromo-3,7-dichlorodibenzo-*p*-dioxin (9.35).

Previous studies reported that models computed by PM3 Hamiltonian suggested a better predictive capability for the toxicity of persistent contaminants than models developed using other semi-empirical methods such as AM1 and MNDO (Niu et al. 2007). Therefore, molecular structural descriptors were calculated for PCDD/Fs by semi-empirical PM3 methods. All calculations were performed using MOPAC (2000) contained in the CS Chem3D Ultra (Version. 6.0). The molecular structures were optimized using eigenvector following, a geometry optimization procedure within MOPAC 2000. The geometry optimization criterion GNORM was set at 0.1.

A total of 22 MOPAC-derived descriptors that reflect the overall characters of the chemicals were computed. A full list is given in Table 2. The energy of intermolecular forces is closely related to M_W values. The parameter α indicates the ease with which the species can be deformed by an electric field. Atomic charges are related to the reactive centers. Two non-empirical descriptors, ΔH_f (heat of molecular formation) and μ (molecular dipole moment), are expected to reflect the affinity for leaching to some extent. Additionally, three combinations of frontier molecular-orbital energies, $E_{LUMO} - E_{HOMO}$, $(E_{LUMO} - E_{HOMO})^2$ and $E_{LUMO} + E_{HOMO}$, which proved to be significant in previous studies (Chen et al. 2001a, b), were also selected as predictor variables. The $E_{LUMO} - E_{HOMO}$ and $E_{LUMO} + E_{HOMO}$ can be related to absolute hardness and electronegativity, respectively (Pearson 1986; Faucon et al. 1999). IP, which is directly related to the E_{HOMO}, chracterizes the susceptibility of the molecule toward attack by electrophiles.

The Simca (Simca-S Version 6.0, Umetri AB and Erisoft AB) software was used to perform the partial least square analysis. The conditions for the computation were based on the default values of the software. The criterion used to determine the model dimensionality – the number of significant partial least square components – is cross validation (CV). With CV, when the fraction of the total variation of the dependent variables that can be predicted by a component, Q^2, for the whole data set is larger than a significance limit (0.097), the tested partial least square component is considered significant. When the cumulative Q^2 for the extracted components, Q^2_{cum}, is larger than 0.5, the model is considered to have a good prediction ability. Model adequacy was mainly characterized by the number of observations used for model building in the training set, the number of partial least square principal components (A), Q^2_{cum}, the correlation coefficient between experimental

Table 1 Experimental and predicted toxicity values of PCDD/Fs

No.	Compounds	log 1/EC_{50} (obs.)[a]	log 1/EC_{50} (pred.)[b]	SE-pred.[c]	Diff.[d]
	PCDDs				
1	1-Chlorodibenzo-*p*-dioxin	4.00	4.10	±0.38	−0.10
2	2,8-Dichlorodibenzo-*p*-dioxin	5.50	6.79	±0.22	−1.29
3	1,2,4-Trichlorodibenzo-*p*-dioxin	4.89	5.23	±0.21	−0.34
4	2,3,6-Trichlorodibenzo-*p*-dioxin	6.66	6.41	±0.16	0.25
5	2,3,7-Trichlorodibenzo-*p*-dioxin	7.15	7.32	±0.21	−0.17
6	1,2,3,4-Tetrachlorodibenzo-*p*-dioxin	5.89	5.23	±0.20	0.66
7	1,3,7,8-Tetrachlorodibenzo-*p*-dioxin	6.10	7.01	±0.15	−0.91
8	2,3,6,7-Tetrachlorodibenzo-*p*-dioxin	6.80	6.69	±0.15	0.11
9	2,3,7,8-Tetrachlorodibenzo-*p*-dioxin	8.00	7.87	±0.25	0.13
10	1,2,3,4,7-Pentachlorodibenzo-*p*-dioxin	5.19	5.70	±0.20	−0.51
11	1,2,4,7,8-Pentachlorodibenzo-*p*-dioxin	5.96	6.17	±0.17	−0.21
12	1,2,3,4,7,8-Hexachlorodibenzo-*p*-dioxin	6.55	6.21	±0.24	0.34
13	1,2,3,4,6,7,8,9-Octachlorodibenzo-*p*-dioxin	5.00	4.63	±0.40	0.37
14	2-Bromo-3,7,8-trichlorodibenzo-*p*-dioxin	7.94	8.23	±0.22	−0.29
15	2,3-Dibromo-7,8-dichlorodibenzo-*p*-dioxin	8.83	8.19	±0.18	0.64
16	2,8-Dibromo-3,7-dichlorodibenzo-*p*-dioxin	9.35	8.66	±0.21	0.69
17	2-Bromodibenzo-*p*-dioxin	6.53	5.95	±0.28	0.58
18	2,7-Dibromodibenzo-*p*-dioxin	7.81	7.42	±0.22	0.40
19	2,3,7-Tribromodibenzo-*p*-dioxin	8.93	7.85	±0.18	1.08
21	1,3,7,8-Tetrabromodibenzo-*p*-dioxin	8.70	8.65	±0.23	0.05
22	2,3,7,8-Tetrabromodibenzo-*p*-dioxin	8.82	8.96	±0.23	−0.14
23	1,3,7,8,9-Pentabromodibenzo-*p*-dioxin	7.03	7.66	±0.29	−0.63
24	1,2,4,7,8-Pentabromodibenzo-*p*-dioxin	7.77	8.46	±0.26	−0.69
25	1,2,3,7,8-Pentabromodibenzo-*p*-dioxin	8.18	8.17	±0.24	0.01
	PCDFs		log 1/EC_{50} (pred.)[e]		
26	Dibenzofuran	3.00	2.50	±0.40	0.50
27	2-Chlorodibenzofuran	3.55	3.48	±0.30	0.07
28	3-Chlorodibenzofuran	4.38	4.07	±0.25	0.31
29	4-Chlorodibenzofuran	3.00	4.11	±0.31	−1.11
30	2,3-Dichlorodibenzofuran	5.33	5.05	±0.16	0.28
31	2,6-Dichlorodibenzofuran	3.61	4.54	±0.21	−0.93
32	2,8-DiChlorodibenzofuran	3.59	3.80	±0.30	−0.21
33	1,3,6-Trichlorodibenzofuran	5.36	5.93	±0.20	−0.57
34	1,3,8-Trichlorodibenzofuran	4.07	4.94	±0.24	−0.87
35	2,3,4-Trichlorodibenzofuran	4.72	5.51	±0.24	−0.79
36	2,3,8-Trichlorodibenzofuran	6.00	5.49	±0.15	0.51
37	2,6,7-Trichlorodibenzofuran	6.35	5.25	±0.15	1.10
38	1,2,4,8-Tetrachlorodibenzofuran	5.00	4.97	±0.25	0.03
39	1,2,3,6-Tetrachlorodibenzofuran	6.46	6.91	±0.20	−0.45
40	1,3,6,8-Tetrachlorodibenzofuran	6.66	5.99	±0.14	0.67
41	1,2,3,7-Tetrachlorodibenzofuran	6.96	6.91	±0.20	0.05
42	2,3,4,6-Tetrachlorodibenzofuran	6.46	6.49	±0.35	−0.03
43	2,3,6,8-Tetrachlorodibenzofuran	6.66	6.56	±0.16	0.10
44	2,3,4,8-Tetrachlorodibenzofuran	6.70	5.73	±0.15	0.97
45	2,3,4,7-Tetrachlorodibenzofuran	7.60	6.58	±0.21	1.02
46	2,3,7,8-Tetrachlorodibenzofuran	7.39	6.02	±0.17	1.37

Table 1 continued

No.	Compounds	log 1/EC_{50} (obs.)[a]	log 1/EC_{50} (pred.)[b]	SE-pred.[c]	Diff.[d]
47	1,2,4,6,7-Pentachlorodibenzofuran	7.17	6.69	±0.21	0.48
48	1,2,4,7,9-Pentachlorodibenzofuran	4.70	6.39	±0.18	−1.69
49	1,2,3,4,8-Pentachlorodibenzofuran	6.92	6.00	±0.17	0.92
50	1,2,3,7,8-Pentachlorodibenzofuran	7.13	6.50	±0.26	0.64
51	1,2,4,6,8-Pentachlorodibenzofuran	5.51	5.98	±0.14	−0.47
52	1,2,4,7,8-Pentachlorodibenzofuran	5.89	6.23	±0.21	−0.34
53	1,2,3,7,9-Pentachlorodibenzofuran	6.40	6.58	±0.27	−0.18
54	1,3,4,7,8-Pentachlorodibenzofuran	6.70	6.50	±0.20	0.21
55	2,3,4,7,9-Pentachlorodibenzofuran	6.70	6.70	±0.17	0.00
56	2,3,4,7,8-Pentachlorodibenzofuran	7.82	6.96	±0.19	0.87
57	1,2,3,4,7,8-Hexachlorodibenzofuran	6.64	6.95	±0.22	−0.31
58	1,2,3,6,7,8-Hexachlorodibenzofuran	6.57	7.44	±0.24	−0.87
59	1,2,4,6,7,8-Hexachlorodibenzofuran	5.08	6.59	±0.16	−1.51
60	2,3,4,6,7,8-Hexachlorodibenzofuran	7.33	7.08	±0.27	0.25

[a] Data from Ashek et al. (2006)

[b] Predicted log 1/EC_{50} values from model (1)

[c] SE-pred.: standard errors for the predicted log 1/EC_{50} values

[d] Diff.: difference between observed and predicted log 1/EC_{50} values

[e] Predicted log 1/EC_{50} values from model (2)

and predicted values (R) and the general standard error (SE) (Niu and Yu 2004; Niu et al. 2006).

Results and Discussion

Variable importance in the projection (VIP) is a parameter in the partial least square analysis that shows the importance of a variable in a partial least square model. Partial least square analysis with log 1/EC_{50} values of PCDD/Fs as independent variable and the 22 quantum chemical descriptors as independent variables generate many results. The optimal model, which has the largest Q^2_{cum}, largest R and smallest p, was obtained through stepwise culling the model with the smallest VIP value.

Following the procedure described above, optimal QSAR models could be obtained. In this study, the molecular structures of 60 PCDD/Fs are different from each other, for example, the parent molecules of PCDDs and PCDFs are dibenzo-*p*-dioxine and dibenzofuran, respectively. Therefore, the 60 PCDD/Fs are divided into two groups for developing significant QSAR models depending on their parent molecules. Herein, models (1) to (3) were developed for 25 PCDDs, 35 PCDFs, combination of 25 PCDDs and 35 PCDFs, respectively. Based on the unscaled pseudo-regression coefficients of the independent variables and constants transformed from partial least square results, analytical QSAR equations from models (1)–(3) were obtained and are shown in Eqs. 1 to 3:

Model (1) for PCDDs:

$$\begin{aligned}\log 1/EC_{50} = {} & 3.281 - 6.588\times10^{-1}(E_{LUMO}-E_{HOMO})^2 \\ & - 7.377\times10^{-4}\,CCR - 2.465\,E_{HOMO} \\ & + 1.109\times10^{-2}\,M_W - 3.648\,(E_{LUMO}+E_{HOMO}) \\ & + 4.370\,E_{LUMO+1} \qquad (1)\end{aligned}$$

$n = 25$, $A = 3$, $R^2_{X(adj.)(cum)} = 0.990$, $R^2_{Y(adj.)(cum)} = 0.856$, $Eig = 0.363$, $Q^2_{cum} = 0.816$, $R = 0.923$, SE = 0.593

Model (2) for PCDFs:

$$\begin{aligned}\log 1/EC_{50} = {} & -46.961 - 1.343\times10^{-1}(E_{LUMO}-E_{HOMO})^2 \\ & - 2.151(E_{LUMO}-E_{HOMO}) - 22.439\,q_C^- \\ & + 38.800\,q_H^+ - 6.590\times10^{-1}\,\mu_y \\ & - 7.701\,E_{HOMO} - 3.38\times10^{-1}\,\mu_x \qquad (2)\end{aligned}$$

$n = 35$, $A = 2$, $R^2_{X(adj.)(cum)} = 0.608$, $R^2_{Y(adj.)(cum)} = 0.707$, $Eig = 1.089$, $Q^2_{cum} = 0.629$, $R = 0.841$, SE = 0.769

Table 2 List of molecular structural descriptors of PCDD/Fs

Symbols	Description
M_W	Molecular weight
ΔH_f	Standard heat of formation (kcal)
TE	Total energy (eV)
EE	Electronic energy (eV)
CCR	Core-core repulsion energy (eV)
E_{HOMO}	The energy of the highest occupied molecular orbital (eV)
E_{HOMO-1}	The energy of the second highest occupied molecular orbital (eV)
E_{LUMO}	The energy of the lowest unoccupied molecular orbital (eV)
E_{LUMO+1}	The energy of the second lowest unoccupied molecular orbital (eV)
q_{Cl}^{+}	The largest positive atomic charge on a chlorine atom (a.c.u)
q_{H}^{+}	The most positive net atomic charges on a hydrogen atom (a.c.u)
q_{C}^{-}	The largest negative atomic charge on a carbon atom (a.c.u)
q_{O}^{-}	The largest negative atomic charge on a oxygen atom (a.c.u)
μ	Dipole moment (Debye)
μ_x	*X*-axis dipole moment (Debye)
μ_y	*Y*-axis dipole moment (Debye)
μ_z	*Z*-axis dipole moment (Debye)
α	Average molecular polarizability (a.u)
IP	Ionization potential (eV)

Model (3) for PCDD/Fs:

$$\log 1/EC_{50} = -27.89 - 3.269E_{HOMO-1} + 3.325\times10^{-2}\alpha - 2.793\times10^{-1}\mu_x + 24.622q_H^{+} - 7.680\times10^{-2}(E_{LUMO}-E_{HOMO})^2 \quad (3)$$

$n = 60$, $A = 2$, $R^2_{X(adj.)(cum)} = 0.549$, $R^2_{Y(adj.)(cum)} = 0.686$, $Eig = 1.140$, $Q^2_{cum} = 0.603$, $R = 0.829$, SE = 0.870

where $R^2_{X(adj.)(cum)}$ and $R^2_{Y(adj.)(cum)}$ stand for cumulative variance of all the *X*'s and *Y*'s, respectively, explained by all extracted components. *Eig* stands for the eigenvalue, which denotes the importance of the principal components of partial least square. For example, the three partial least square components selected in model (1) explain 99.0% of the variance of the predictor variables, and 85.6% of the variance of the dependent variable.

Models (1)–(3) obtained from partial least square analysis for 25 PCDDs, 35 PCDFs, 25 PCDDs and 35 PCDFs together suggested that molecular structural characteristics of the three groups of PCDD/Fs affect the toxicity of these chemicals. As all the cross-validated Q^2_{cum} values of models (1)–(3) are larger than 0.50, these models are surely stable and have good prediction capability. Compared the Q^2_{cum} values of models (1)–(2) with that of the model (3), it can be concluded that the partial least square models developed for PCDDs and PCDFs, respectively, are more significant than that for PCDD/Fs together. Thus the partial least square models (1) and (2) are more reliable than model (3). This may be reasonable since the parent structures of PCDDs and PCDFs are different from each other.

Table 3 The variable importance in the projection (VIP) and partial least square weights (*W**) values for the molecular structural descriptors included in models (1)–(3)

Models	Variables	VIP	*W**[1]	*W**[2]	*W**[3]
(1)	$(E_{LUMO} - E_{HOMO})^2$	1.591	−0.055	0.007	−0.939
	CCR	1.020	−0.073	−0.866	−0.769
	E_{HOMO}	0.830	−0.533	−0.263	0.047
	M_W	0.801	0.615	0.431	0.392
	$E_{LUMO} + E_{HOMO}$	0.756	−0.534	−0.244	−0.405
	E_{LUMO+1}	0.726	−0.210	0.439	0.444
(2)	$(E_{LUMO} - E_{HOMO})^2$	1.349	−0.557	−0.177	
	$E_{LUMO} - E_{HOMO}$	1.343	−0.554	−0.169	
	q_C^-	0.994	−0.409	−0.039	
	q_H^+	0.987	0.404	−0.020	
	μ_y	0.748	−0.018	−0.706	
	E_{HOMO}	0.658	−0.216	−0.432	
	μ_x	0.649	0.073	−0.568	
(3)	E_{HOMO-1}	1.338	−0.620	−0.552	
	α	1.273	0.603	0.077	
	μ_x	0.753	0.186	−0.814	
	q_H^+	0.750	0.350	0.279	
	$(E_{LUMO} - E_{HOMO})^2$	0.679	−0.309	0.171	

Models (1)–(3) include six, seven and five predictor variables, respectively. VIP values and partial least square weights for the variables are listed in Table 3. In model (1), the VIP values of $(E_{LUMO} - E_{HOMO})^2$ and CCR are higher than 1.0, while the VIP values of E_{HOMO}, M_W, $E_{LUMO} + E_{HOMO}$ and E_{LUMO+1} are lower than 1.0. The results show the former two descriptors are relative more important than the latter four variables. Interpretation of each principal component of partial least square is based on the respective magnitudes of the loadings assigned to the variables. We can see how much a single variable in each partial least square component contributes to modeling.

In model (1), the first component of partial least square is mainly related to E_{HOMO}, $E_{LUMO} + E_{HOMO}$, and M_W, of which the absolute *W**[1] values are higher than those of other descriptors. The second component of partial least square is related to CCR and the third component is related to $(E_{LUMO} - E_{HOMO})^2$ and CCR. As indicated by Eq. 1,

increasing M_W and E_{LUMO+1} leads to increasing log $1/EC_{50}$ values. On the contrary, increasing $(E_{LUMO} - E_{HOMO})^2$, CCR, E_{HOMO} and $E_{LUMO} + E_{HOMO}$ leads to decreasing log $1/EC_{50}$ values. According to Pearson (Pearson 1986), absolute hardness can be defined as $1/2(E_{LUMO} - E_{HOMO})$, which is regarded as a measure of energy stabilization in chemical systems and chemical structures tend be more stable at larger values of the $E_{LUMO} - E_{HOMO}$ gap (Niu et al. 2005; Faucon et al. 1999). It thus can be concluded that PCDD/Fs with big absolute hardness values tend to be more stable and suggest lower toxicity values. Thus it is reasonable that increasing $(E_{LUMO} - E_{HOMO})^2$ or $E_{LUMO} - E_{HOMO}$ leads to decreasing toxicities.

In model (2), descriptors with VIP values larger than 1.0 include $(E_{LUMO} - E_{HOMO})^2$ and $E_{LUMO} - E_{HOMO}$, which suggests that they are more important than other variables. Table 3 shows that the first partial least square component is mainly related to $(E_{LUMO} - E_{HOMO})^2$ and $E_{LUMO} - E_{HOMO}$. The second component of partial least square is mainly related to μ_x. Equation 2 suggests that increasing q_H^+ values leads to an increase of log $1/EC_{50}$ values, while increasing $(E_{LUMO} - E_{HOMO})^2$, $E_{LUMO} - E_{HOMO}$, q_C^-, μ_y, E_{LUMO} and μ_x results in a deline of log $1/EC_{50}$ values. μ is used to describe the polarity of the molecule. PCDD/Fs with great μ values may be harder to get into the body of organisms, resulting in low toxicity values.

In model (3), the first component of partial least square is mainly related to E_{HOMO-1} and α, and the second component is related to μ_x and E_{HOMO-1}. As shown in Eq. 3, the higher α and q_H^+ values, the higher are the log $1/EC_{50}$ values. However, higher values of E_{HOMO-1}, μ_x and $(E_{LUMO} - E_{HOMO})^2$ lead to a decrease of log $1/EC_{50}$ values. Previous experimental and theoretical studies has indicated that the interaction between halogenated aromatic hydrocarbons and the aryl hydrocarbon receptors is a charge transfer process and that PCDD/F molecules act as electron acceptors in the charge transfer complex (Poland et al. 1976; Miller et al. 1977; Cheney and Tolly 1979). Increasing ability of PCDD/Fs to accept electrons leads to larger binding affinity values for aryl hydrocarbon receptors and higher toxicity values. E_{HOMO} is directly related to the ionization potential and characterizes the susceptibility of the molecule toward attack by electrophiles. Molecules with high E_{HOMO} values tend to easily lose electrons. And E_{HOMO-1} also describes this kind of property to some extent. So it can be concluded that the higher the E_{HOMO-1} values, the poorer binding affinity for aryl hydrocarbon receptors and lower toxicities. The descriptor α characterizes the molecular polarizability of chemicals, and molecules with great α values may have great intermolecular dispersive forces with aryl hydrocarbon receptor molecules, causing high log $1/EC_{50}$ values. q_H^+ is the most positive net atomic charges on a hydrogen atom. PCDD/Fs with great q_H^+ values may have great intermolecular electrostatic attractive interaction with aryl hydrocarbon receptor molecules, resulting in large toxicities.

Acknowledgments This work was supported by the National Key Technology R&D Program of China (No. 2008BAC32B05-2).

References

Ashek A, Lee C, Park H, Cho SJ (2006) 3D QSAR studies of dioxins and dioxin-like compounds using CoMFA and CoMSIA. Chemosphere 65:521–529

Bayen S, Lee HK, Obbard JP (2007) Exposure and response of aquacultured oysters, *Crassostrea gigas*, to marine contaminants. Environ Res 103:375–382

Bjorn B, Krister H (2000) Ecotoxicological risk assessment of environmental pollutants in the Arctic. Toxicol Lett 112:111–118

Bustnes JO, Helberg M, Strann K, Skaare JU (2006) Environmental pollutants in endangered vs. increasing subspecies of the lesser black-backed gull on the Norwegian Coast. Environ Pollut 144:893–901

Chen JW, Peijnenburg WJGM, Quan X, Chen S, Martens D, Schramm K-W, Kettrup A (2001a) Is it possible to develop a QSPR model for direct photolysis half-lives of PAHs under irradiation of sunlight? Environ Pollut 114:137–143

Chen JW, Quan X, Yan Y, Yang FL, Peijnenburg WJGM (2001b) Quantitative structure–property relationship studies on direct photolysis of selected polycyclic aromatic hydrocarbons in atmospheric aerosol. Chemosphere 42:263–270

Cheney BV, Tolly T (1979) Electronic factors affecting receptor binding of dibenzo-*p*-dioxins and dibenzofurans. Int J Quantum Chem 16:87–110

Cheung ENY, McKinney JD (1989) Polybrominated naphthalene and diiodobenzene interactions with specific binding sites for 2,3,7,8-tetrachlorodibenzo-*p*-dioxin in rat liver cytosol. Mol Toxicol 2:39–52

Clare BW (2006) QSAR of aromatic substances: toxicity of polychlorodibenzofurans. J Mol Struct 763:205–213

Devillers J, Marchand-Geneste N, Porcher JM (2006) SAR and QSAR modeling of endocrine disruptors. SAR QSAR Environ Res 17:393–412

Faucon JC, Bureau R, Faisant J, Briens F, Rault S (1999) Prediction of the fish acute toxicity from heterogeneous data coming from notification files. Chemosphere 38:3261–3276

Hu JY, Aizawa T (2003) Quantitative structure–activity relationships for estrogen receptor binding affinity of phenolic chemicals. Water Res 37:1213–1222

Huang XY, Lessner L, Carpenter DO (2006) Exposure to persistent organic pollutants and hypertensive disease. Environ Res 102: 101–106

Juranić IO, Drakulić BJ, Petrović SD, Mijin DZ, Stanković MV (2006) A QSAR study of acute toxicity of *N*-substituted fluoroacetamides to rats. Chemosphere 62:641–649

Long XX, Niu JF (2007) Estimation of gas-phase reaction rate constants of alkylnaphthalenes with chlorine, hydroxyl and nitrate radicals. Chemosphere 67:2028–2034

Miller G, Sontum S, Crosby DG (1977) Electron-acceptor properties of chlorinated dibenzo-*p*-dioxins. Environ Contam Toxicol 18: 611–616

Netzeva TI, Schultz TW (2005) QSARs for the aquatic toxicity of aromatic aldehydes from Tetrahymena data. Chemosphere 61:1632–1643

Niu JF, Yu G (2004) Prediction of the ability of PAHs to be photocytotoxic to a cell line from the rainbow trout (*Oncorhynchus mykiss*) gill. Bull Environ Contam Toxicol 73:659–665

Niu JF, Chen JW, Henkelmann B, Quan X, Yang FL, Kettrup A, Schramm K-W (2003) Photodegradation of PCDD/Fs adsorbed on spruce (*Picea abies* (L.) Karst.) needles under sunlight irradiation. Chemosphere 50:1217–1225

Niu JF, Huang LP, Chen JW, Yu G, Schramm K-W (2005) Quantitative structure–property relationships on photolysis of PCDD/Fs adsorbed to spruce (*Picea abies* (L.) Karst.) needle surfaces under sunlight irradiation. Chemosphere 58:917–924

Niu JF, Shen ZY, Yang ZF, Long XX, Yu G (2006) Quantitative structure–property relationships on photodegradation of polybrominated diphenyl ethers. Chemosphere 64:658–665

Niu JF, Long XX, Shi SQ (2007) Quantitative structure–activity relationships for prediction of the toxicity of hydroxylated and quinoid PCB metabolites. J Mol Model 13:163–169

Okey AB (1990) Enzyme-induction in the cytochrome-p-450 system. Pharmacol Ther 45:241–298

Padmanabhan J, Parthasarathi R, Subramanian V, Chattaraj PK (2006) QSPR models for polychlorinated biphenyls: *n*-octanol/water partition coefficient. Med Chem 14:1021–1028

Pearson PG (1986) Absolute electronegativity and hardness correlated with molecular orbital theory. Proc Natl Acad Sci 83:8440–8441

Poland A, Glover E, Kende AS (1976) Stereospecific, high affinity binding of 2,3,7,8-tetrachlorodibenzo-para-dioxin by hepatic cytosol-evidence that binding species is receptor for induction of aryl-hydrocarbon hydroxylase. J Biol Chem 251:4936–4946

Waller CL, McKinney JD (1992) Comparative molecular-field analysis of polyhalogenated dibenzo-*p*-dioxins, dibenzofurans, and biphenyls. J Med Chem 35:3660–3666

Wold S, Sjöström M, Ericksson L (2001) PLS-regression: a basic tool of chemometrics. Chemomet Intell Lab Syst 58:109–130

Zhang ZY, Niu JF, Zhi X (2008) A QSAR model for predicting mutagenicity of nitronaphthalenes and methylnitronaphthalenes. Bull Environ Contam Toxicol 81:498–502